SQL QUERIES

```
/* RENAME COLUMNS */

EXEC SP_RENAME 'UBER.COLUMN1', 'STARTDATETIME', 'COLUMN';

EXEC SP_RENAME 'UBER.COLUMN2', 'ENDDATETIME', 'COLUMN';

EXEC SP_RENAME 'UBER.COLUMN3', 'CATEGORY', 'COLUMN';

EXEC SP_RENAME 'UBER.COLUMN4', 'START_LOCATION', 'COLUMN';

EXEC SP_RENAME 'UBER.COLUMN5', 'END_LOCATION', 'COLUMN';

EXEC SP_RENAME 'UBER.COLUMN6', 'MILES', 'COLUMN';

EXEC SP_RENAME 'UBER.COLUMN7', 'PURPOSE', 'COLUMN';


/* EXPLORING AND CLEANING PURPOSE COLUMN */

UPDATE UBER

SET PURPOSE = 'OTHER'

WHERE PURPOSE IS NULL

UPDATE UBER

SET PURPOSE = REPLACE(PURPOSE, 'CHARITY ($)','CHARITY')


/* DELETED THE LAST ROW CONTAINING UNWANTED NULL VALUES*/

WITH CTE AS(

SELECT ROW_NUMBER() OVER(ORDER BY CATEGORY) AS RNUM,* FROM UBER

)

DELETE FROM CTE WHERE RNUM = 1


/*CORRECTING WRONG ENTRIES(SPELLING MISTAKE)*/

UPDATE UBER

SET START_LOCATION = REPLACE(START_LOCATION, 'KAR?CHI','KARACHI')

UPDATE UBER

SET START_LOCATION = REPLACE(START_LOCATION, 'R?WALPINDI','RAWALPINDI')

UPDATE UBER
```

```
SET END_LOCATION = REPLACE(END_LOCATION, 'KAR?CHI','KARACHI')

UPDATE UBER

SET END_LOCATION = REPLACE(END_LOCATION, 'R?WALPINDI','RAWALPINDI')


/* ADDING HOUR COLUMN*/

ALTER TABLE UBER

ADD HRS INT

UPDATE UBER

SET HRS = DATEPART(HOUR, STARTDATETIME)


/*LIMITING THE MILES VALUE TO 2 DECIMAL POINTS*/

UPDATE UBER

SET MILES = CAST(MILES AS DECIMAL(10,2))


/* TIME DIFFERENCE*/

ALTER TABLE UBER

ADD TIMETAKEN INTEGER

UPDATE UBER SET TIMETAKEN = DATEDIFF(MINUTE, STARTDATETIME, ENDDATETIME)


/* DATE COLUMN ADDED*/

ALTER TABLE UBER

ADD DTE DATE

UPDATE UBER

SET DTE = CAST(STARTDATETIME AS DATE)


/* DELETING ABNORMAL ROWS/ DATA HAVING NON-NULL MILES VALUE WITH TIME TAKEN AS JUST 0,

THAT'S NOT POSSIBLE*/

DELETE  FROM UBER WHERE START_LOCATION = 'UNKNOWN LOCATION' AND END_LOCATION =

'UNKNOWN LOCATION'
```

```
SELECT * FROM UBER WHERE  START_LOCATION = 'KARACHI' AND END_LOCATION = 'KARACHI'
ORDER BY TIMETAKEN ASC
SELECT * FROM UBER WHERE  START_LOCATION = 'ISLAMABAD' AND END_LOCATION = 'ISLAMABAD'
ORDER BY TIMETAKEN ASC
SELECT * FROM UBER WHERE START_LOCATION = 'UNKNOWN LOCATION' AND END_LOCATION =
'UNKNOWN LOCATION' ORDER BY MILES ASC
DELETE FROM UBER WHERE  START_LOCATION = 'UNKNOWN LOCATION' AND END_LOCATION =
'UNKNOWN LOCATION' AND PURPOSE = 'TEMPORARY SITE' AND MILES = 12.3
AND TIMETAKEN = 9 AND CATEGORY = 'BUSINESS' AND STARTDATETIME = '2016-12-22 18:38:00.0000000'
AND ENDDATETIME = '2016-12-22 18:47:00.0000000'
DELETE FROM UBER WHERE  START_LOCATION = 'KARACHI' AND END_LOCATION = 'KARACHI' AND
PURPOSE = 'OTHER' AND MILES = 3.6
AND TIMETAKEN = 0 AND CATEGORY = 'BUSINESS' AND STARTDATETIME = '2016-10-08 15:03:00.0000000'
AND ENDDATETIME = '2016-10-08 15:03:00.0000000'
DELETE FROM UBER WHERE  START_LOCATION = 'ISLAMABAD' AND END_LOCATION = 'ISLAMABAD'
AND PURPOSE = 'OTHER' AND MILES = 0.7
AND TIMETAKEN = 0 AND CATEGORY = 'BUSINESS' AND STARTDATETIME = '2016-10-13 13:02:00.0000000'
AND ENDDATETIME = '2016-10-13 13:02:00.0000000'



/* CREATING TABLES*/
/*MONTHY OR SEASONAL */
ALTER TABLE UBER
DROP COLUMN MNTH
UPDATE UBER
SET MNTH = DATEPART(MONTH, DTE)
SELECT * FROM UBER
SELECT DISTINCT MNTH FROM UBER ORDER BY MNTH
ALTER TABLE UBER
ADD MNTH_NAME VARCHAR(30)
```

```
UPDATE UBER
SET MNTH_NAME = CASE
                WHEN MNTH = 1 THEN 'JANUARY'
                WHEN MNTH = 2 THEN 'FEBRUARY'
                WHEN MNTH = 3 THEN 'MARCH'
                WHEN MNTH = 4 THEN 'APRIL'
                WHEN MNTH = 5 THEN 'MAY'
                WHEN MNTH = 6 THEN 'JUNE'
                WHEN MNTH = 7 THEN 'JULY'
                WHEN MNTH = 8 THEN 'AUGUST'
                WHEN MNTH = 9 THEN 'SEPTEMBER'
                WHEN MNTH = 10 THEN 'OCTOBER'
                WHEN MNTH = 11 THEN 'NOVEMBER'
                WHEN MNTH = 12 THEN 'DECEMBER'
                ELSE 'UNKNOWN'
            END;


/* MONTHWISE -FRQUENCY*/
SELECT MNTH_NAME, COUNT(*) AS CNT FROM UBER GROUP BY MNTH_NAME ORDER BY CNT DESC
SELECT MNTH_NAME, CATEGORY, COUNT(*) AS CNT FROM UBER  GROUP BY MNTH_NAME,CATEGORY
ORDER BY MNTH_NAME, CATEGORY, CNT ASC

/* PURPOSE BASED*/
SELECT MIN(CAST(MILES AS DECIMAL(10,2))) AS AVG_MILES
FROM UBER
GROUP BY PURPOSE
ORDER BY RIDE_COUNT DESC;
```

```
/* FREQUENTLY USED ROUTES*/

SELECT TOP 5 START_LOCATION, END_LOCATION, CATEGORY, COUNT(*) AS FREQUENCY FROM UBER

GROUP BY START_LOCATION, END_LOCATION,  CATEGORY

ORDER BY FREQUENCY  DESC, CATEGORY

SELECT COUNT(*) AS CNT FROM UBER GROUP BY START_LOCATION, END_LOCATION

ORDER BY CNT DESC


/*CALCULATING MINIMUM, MAXIMIUM AND AVERAGE TIME TAKEN TO TRAVEL*/

SELECT MAX(TIMETAKEN) AS MAXIMUM_TIME, MIN(TIMETAKEN) AS MINIMUM_TIME,

AVG(TIMETAKEN) AS AVERAGE_TIME FROM UBER
```